Computers & Industrial Engineering 128 (2019) 911–919

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

# Modelling food sourcing decisions under climate change: A data-driven approach

Rengarajan Srinivasan[a,*], Vaggelis Giannikas[b], Mukesh Kumar[a], Renaud Guyot[a], Duncan McFarlane[a]

[a] Institute for Manufacturing, University of Cambridge, 17 Charles Babbage Road, Cambridge CB3 0FS, United Kingdom
[b] School of Management, University of Bath, Claverton Down, Bath BA2 7AY, United Kingdom

ARTICLE INFO

*Keywords:*
Climate change
Food supply chain
Open data
Sourcing decision
Sourcing risk

ABSTRACT

Changes in climate conditions are expected to pose significant challenges to the food industry, as it is very likely that they will affect the production of various crops. As a consequence, decisions associated with the sourcing of food items will need to be reconsidered in the years to come. In this paper, we investigate how environmental changes are likely to affect the suitability and risk of different regions—in terms of growing certain food items—and whether companies should adapt their sourcing decisions due to these changes. In particular, we propose a three-stage approach that guides food sourcing decisions by incorporating climate change data. The methodology utilises environmental data from several publicly available databases and models weather uncertainties to calculate the suitability and risk indices associated with growing a crop in a particular geographical area. The estimated suitability and risk parameters are used in a mean-variance analysis to calculate the optimal sourcing decision. Results from a case example indicate that sourcing decisions of popular food items are likely to require significant adaptations due to changes to the suitability of certain regions.

## 1. Introduction

Climate change is expected to impact on the distribution of fresh water resources, global biogeochemical cycles, agricultural output, and ice sheets (Scheffer, Carpenter, Foley, Folke, & Walker, 2001) which in turn will have a huge impact on food production and food supply chains (Sussman & Freed, 2008). There is a consensus that climate change will present new challenges to supply chains and business regarding possible regulations for greenhouse gas emissions, changing attitudes, and evolving markets (Cogan, 2006; Mills & Lecomte, 2006). Also, the physical effects of climate change—shifts in temperature and precipitation patterns, the prevalence of weather extremes, changes in sea level—will impact on operational procedures, physical and human assets, and access to resources (Patwardhan et al., 2007). In addition, managing unexpected changes in the external environment is often considered one of the most difficult tasks facing organisations (Linnenluecke & Griffiths, 2010). Therefore climate change has become one of the important factors that have to be evaluated for food sourcing decisions.

Multi-national corporations mostly source their food products from tier 1 suppliers, who may be involved in crop production. As climate is changing, these corporations need to evaluate these suppliers within the context of climate change. There are fundamental questions faced by these organisations:

- Existing supplier selection model incorporates economics, social and environmental sustainability criterias but lacks long-term climate change factors (Snyder et al., 2016).
- Dual and multi sourcing strategies need to incorporate climate change (Heese, 2015).
- Crop production locations with climate change impact need to be considered by Tier 1 suppliers.
- Pricing strategies for the food sourced should include climate change factors (Kumar, Basu, & Avittathur, 2018).

There are studies that suggest there will be food shortages and price fluctuations of key raw materials. For example, in the United Kingdom, the iceberg lettuce shortage in 2017 was attributed to changes in climate (BBC, 2017; HBS, 2017). Therefore, traditional models of food sourcing decisions need to be updated, where huge data in multiple

* Corresponding author.
*E-mail addresses:* rs538@eng.cam.ac.uk (R. Srinivasan), v.giannikas@bath.ac.uk (V. Giannikas), mk501@eng.cam.ac.uk (M. Kumar), rg521@eng.cam.ac.uk (R. Guyot), dcm@eng.cam.ac.uk (D. McFarlane).

https://doi.org/10.1016/j.cie.2018.10.048
Received 31 January 2018; Received in revised form 28 July 2018; Accepted 25 October 2018
Available online 26 October 2018
0360-8352/ © 2018 Elsevier Ltd. All rights reserved.

formats on climate change, and geographical dispersed crop production information involving localised crop suitability, needs to be analysed together. Hence, companies' supplier selection models, location decisions, dual sourcing strategies and pricing will need to incorporate big data analytics from the perspective of climate change.

In this paper, we use publicly available datasets such as WorldClim[1] for climate change and FAO Ecocrop database[2] for crop suitability to model food sourcing decisions. The approach proposed in this paper is a big data driven approach as it utilises data of high volume and of a certain variety. We also produce our own simulated data in the second stage of our approach. The approach itself is flexible to use data of greater size and from more data sources a greater level of granularity is required. We argue that sourcing decision modelling has reached to maturity, however emerging risk of crop production suitability arising from climate change is not integrated with existing food sourcing decision models. Hence, this research is one of the first attempts to address how companies can make food sourcing decisions within the context of climate change. A modelling approach is developed to address this challenge by involving data, data sources, uncertainty assessment within data sources and integration of climate change data with sourcing model. This research aims to prepare organisations for climate change adaptation by providing an approach and data analytic tool for procuring food in a changing climate environment.

The paper is structured as follows: Section 2 presents the relevant literature. Section 3 describes the analytic approach for the sourcing decision by integrating climate change data. Section 4 illustrates the proposed approach with a case example exploring the sourcing decisions of banana, cocoa beans and barley. Finally, Section 5 presents the conclusions and future work.

## 2. Literature review

Currently, supply chains are complex and require careful planning to mitigate disruptions. One strategy is to build flexibility from the supply side by developing multiple options for sourcing. To mitigate supply uncertainty, companies can place order with multiple suppliers simultaneously (routine sourcing) or have contingent suppliers to use during disruptions (contingent sourcing) (Snyder et al., 2016). Most sourcing problems in the literature consider a single buyer with several suppliers. The solutions consists of selecting the number of suppliers, which suppliers to choose and how to allocate the orders among the different suppliers (Snyder et al., 2016).

For instance Berger, Gerstenfeld, and Zeng (2004), Meena, Sarmah, and Sarkar (2011) and Ruiz-Torres and Mahmoodi (2007) have considered the problem of selecting the optimal number of suppliers by modelling the cost as a function of number of suppliers or as purchasing and management costs. Others have consider non-identical suppliers and used various metrics such as supplier dependent variable cost, stock out penalty (Dada, Petruzzi, & Schwarz, 2007; Federgruen & Yang, 2008), supplier dependent fixed cost, service level costs to determine the optimal number of suppliers. The problem of allocating orders to different suppliers have been studied by various authors (Sawik, 2011; Yu, Zeng, & Zhao, 2009).

Many different modelling approaches have been proposed in the literature to model the sourcing decision as a supplier selection problem. In order to include qualitative and quantitative goals, the problem is generally modelled as a multi-objective optimisation problem. Different modelling techniques have been used such as mathematical programming, game theoretic and genetic algorithms (Snyder et al., 2016; Wu & Barnes, 2011). Other researchers have considered the combined problem of supplier selection and order allocation. For instance, Demirtas and Ustun propose a two stage modelling approach, where in the first stage an analytic network process is used to model supplier selection and in the second stage, mixed integer linear programming is used to optimise the order allocation between the chosen suppliers (Demirtas & Üstün, 2008). It has been noted that despite abundant literature on supplier selection models, there is a lack of quantitative approaches for building a resilient portfolio of suppliers, especially incorporating climate change as a key variable (Hammami, Temponi, & Frein, 2014).

Recently, there has been considerable interest in big data to make operational decisions. Big data research focuses on holistic approach to data analytics to create actionable insight for companies and governments (Wamba, Akter, Edwards, Chopin, & Gnanzou, 2015). Big data dimension includes 5Vs- volume, velocity, variety, veracity and value (Mehmood, Meriton, Graham, Hennelly, & Kumar, 2017; Wamba et al., 2015). The definition of these dimensions of big data are still evolving (Gandomi & Haider, 2015). It has been noted that the concept of big data is still seen as abstract and industrially, there are different opinions on its definition relative to the nature of an organisation's operations (Addo-Tenkorang & Helo, 2016). Various research focuses on big data applications in multiple sectors such as agriculture and manufacturing making the dimension of big data context dependant (Addo-Tenkorang & Helo, 2016; Wolfert, Ge, Verdouw, & Bogaardt, 2017; Zhong, Newman, Huang, & Lan, 2016). Big data technologies for storage, data processing, visualisation and analytics have received considerable attention (Zhong et al., 2016). Though the importance of big data has been highlighted by various researchers, their specific application in the context of supply chain has not been investigated thoroughly (Gunasekaran et al., 2017; Tiwari, Wee, & Daryanto, 2018).

Additionally, there is drastic rise in publicly available data in the interest of fostering open innovation. These available datasets come from disparate sources with varying data quality. There is a lack of understanding in how these publicly available big data can be used by organisations to improve operational performance. In this research, we explore the application of publicly available big data from WorldClim and FAO Ecocrop to investigate the impact of climate change on food sourcing decisions.

## 3. Model development

### 3.1. Overview

In this section we propose a modelling approach to evaluate the impact of climate change on sourcing decisions. The rationale behind this approach can be summarised in the following. As climate changes, the bioclimatic conditions that impact the growth of crops (such as temperature and rainfall) will also change. As a result, different regions of the planet are expected to become more or less suitable and risky for the growth of certain crops. The change in the suitability of a region to grow a crop will in turn impact the sourcing decisions of a company which aims to minimise the cost of its operations.

The approach has three sequential stages and is graphically illustrated in Fig. 1.

The approach is described in more detail and stage by stage in the following sections. Where possible, we differentiate between the aspects of the approach that are generic and the ones we use as an illustration example in this paper.

### 3.2. Stage 1: Modelling weather uncertainties

In the first stage, weather uncertainties are modelled based on available historical data. By *weather uncertainties* we refer to how the actual realisations of the monthly atmospheric variables spread around the mean value published in climate models. For example, a weather prediction model might predict that the minimum temperature in January 2080 in New York City will be $-2\,^{\circ}C$ but the actual minimum temperature might be between $-3.5\,^{\circ}C$ and $-0.5\,^{\circ}C$ following a certain

[1] http://www.worldclim.org/.
[2] http://ecocrop.fao.org/ecocrop/srv/en/home.

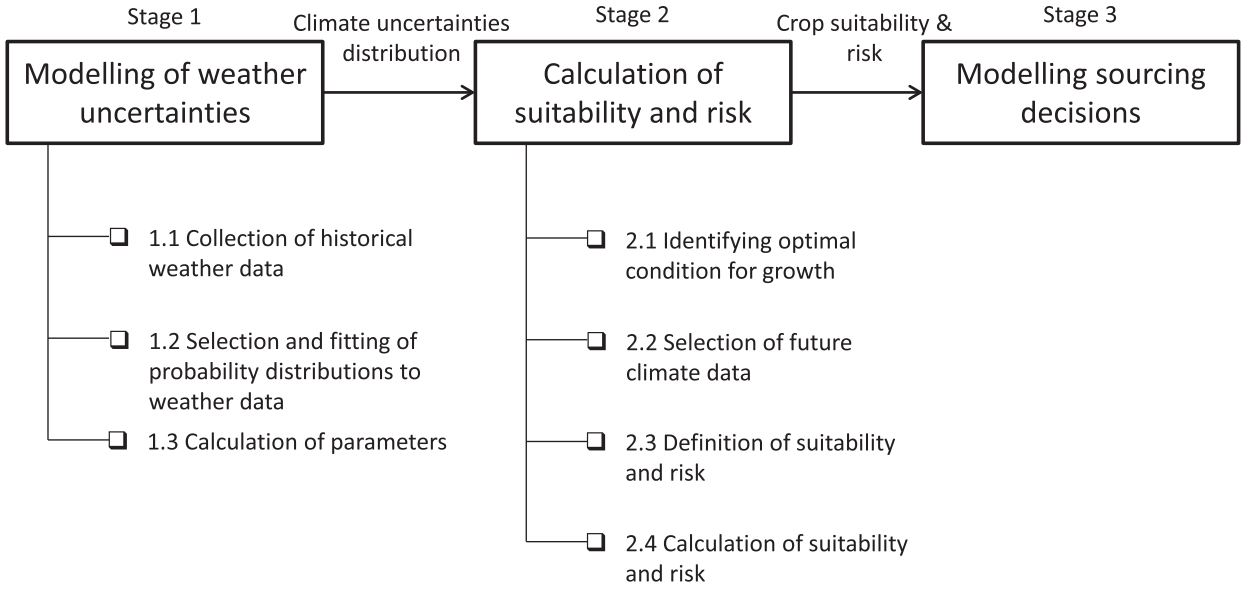
Stage 1
Modelling of weather uncertainties
Climate uncertainties distribution
Stage 2
Calculation of suitability and risk
Crop suitability & risk
Stage 3
Modelling sourcing decisions
1.1 Collection of historical weather data
1.2 Selection and fitting of probability distributions to weather data
1.3 Calculation of parameters
2.1 Identifying optimal condition for growth
2.2 Selection of future climate data
2.3 Definition of suitability and risk
2.4 Calculation of suitability and risk

**Fig. 1.** Modelling approach to assess the impact of climate change on sourcing decisions.

distribution. This is done by analysing historical data (of actual values of atmospheric variables) and then fitting distributions that properly describe them. An assumption is made that the distribution describing weather uncertainties will remain the same in the future but the parameters describing the distribution (e.g. mean and standard deviation) will change. The rationale behind this assumption is that even if climate changes the physics defining the spread around the mean values will remain the same over time. As a consequence, weather uncertainty distributions obtained can be extrapolated to future situations.

There are three sequential steps in Stage 1.

*Step 1.1: Collection of historical weather data*

In this step, necessary historical weather data, which will enable us to model weather uncertainties, are collected. Data for three atmospheric variables are collected: maximum temperature per month, minimum temperature per month and average rainfall per month.

As weather conditions vary between geographical regions, it is important to collect historical data from weather stations in areas that can accurately describe the region of study. In this paper, as an example, we study the entire region and we select four cities geographically placed in three different continents: Beijing, London, Marseille and New York City. Table 1 shows the data range and source for each of the chosen cities.

While using the approach, the interested researcher can define the specific region under consideration, choose multiple weather stations and identify data sources that can be used to reduce location biases. One of the challenges in performing this step is the limited availability of data for specific regions. For instance, gathering publicly available historical data for regions in Africa, south east Asia and south America can be challenging. Another challenge relates to the size of available data. Even if using a very large amount of data can increase the reliability of fitting distributions to the data (in the next step), the actual computing time required might be very long.

**Table 1**
Historical data to model weather uncertainty.

| City | Begin date | End date | Duration | Source |
|---|---|---|---|---|
| Beijing | 1940 | 2004 | 64 years | NDNC |
| London | 1948 | 2015 | 67 years | NDNC |
| Marseille | 1951 | 2015 | 64 years | Klein Tank |
| New York | 1948 | 2015 | 67 years | NDNC |

*Step 1.2: Selection and fitting probability distributions to weather data*

In this step, suitable probability distributions that can represent the historical weather data are selected and fit. The distributions are used to describe weather patterns that we refer to as weather uncertainties. In this study, the normal distribution is chosen to represent the variations in temperature and the gamma distribution to represent the variations in rainfall as rainfall cannot take negative values. This choice is due to the fact that most natural phenomena follow normal-type distributions.

Normal and gamma distributions can be described by their mean and variance. Climate models publish the mean values of atmospheric variables but give no or limited indication regarding the deviation from the mean. Thus, in order to be able to use the distributions describing weather uncertainties (derived using historical data), we need to define the variance as a factor of the variances of each month. We propose eight different options that could be used to derive the appropriate variance, shown in Table 2. An annual adjustment is included in some options in order to acknowledge the potential correlation between variances of different months. The annual adjustment scales the sum of the monthly rainfall values or the mean temperature to match the annual value.

In order the select the best fit parameters, chi-square tests are performed and the resulting p values are noted for each the defined models. A scoring system is defined to select the best fit distribution model (see appendix for the scoring system and the resulting scores for the defined models).

Based on the analysis of the historical data described in the previous step, the distribution that best describes temperature uncertainties is a normal distribution whose variance is equal to the monthly variances of the cities. Similarly, for the rainfall, the distribution that best describes the data is a gamma distribution with a variance that is a scaling coefficient such that the sum of the monthly variances equals the annual variance. In this study, we specifically focussed on the historical data of the four cities (Table 1), which in general models the short to medium term trends. However, as new data are added, it is possible to select the most appropriate probability distribution that represents the data.

*Step 1.3: Calculation of parameters*

The aim of this step is to calculate the bio-climatic variables from historical data and use them to represent the weather uncertainty. As the historical data used in this stage (i.e. temperature min and max and precipitation), do not take into account trends associated with different

**Table 2**
Choice of different variance types that is used to fit the weather data.

| Model Number | Variance | Annual adjustment |
|---|---|---|
| 1 | Equal for every month such as the sum of the monthly variances equals the annual variance | No |
| 2 | Percentage of mean such that the sum of monthly variances equals the annual variance | No |
| 3 | Minimum of all months | No |
| 4 | Mean of all months | No |
| 5 | Maximum of all months | No |
| 6 | Minimum of all months | Yes |
| 7 | Mean of all months | Yes |
| 8 | Maximum of all months | Yes |

time of the year (such as seasonality), it is important to convert these data in variables established by the climate community which are generally used to study climate behaviour. These variables are known as *bio-climatic variables* and there are 19 of them in total (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005). The use of bio-climatic variables will not only allow us to better estimate climate behaviour but also to use the distributions fit in this stage, in the following stage to compare current and future climate scenarios. Thus, in this step we develop a mechanism to represent the weather uncertainty from historical data as a function of bio-climatic variables.

In this study we have chosen bio3 (isothermality) and bio4 (temperature seasonality) to represent the temperature. To represent rainfall, we use the difference of bio13 (precipitation of wettest month) and bio14 (precipitation of driest month), along with bio15 (precipitation seasonality). These variables are chosen because they can describe the extremities and the seasonality in a weather pattern. In general, additional bio-climatic variables can be defined to increase the precision of the overall approach. The values of these bio-climatic variables using the historical data are shown in Table 3.

Next, using the calculated bio-climatic variables, we estimate the parameters of the probability distribution. Linear least square estimation is used to fit the bio-climatic variables to the parameters. For representing temperature uncertainty from historical data, the linear fit yields:

$$\text{Std dev Annual } T_{min} = 0.15\ \text{bio3} + 0.001\ \text{bio4} - 4.53$$

$$\text{Std dev Annual } T_{max} = 0.15\ \text{bio3} + 0.001\ \text{bio4} - 7.21$$

$$\text{Mean Std dev } T_{min} = 1.04\ \text{Annual std dev } T_{min} + 0.738$$

$$\text{Mean Std dev } T_{max} = 1.04\ \text{Annual std dev } T_{max} + 0.90$$

Similarly for the rainfall, which is a gamma distribution (from the previous step), the fit to the data yields the following:

$$\text{Std dev of annual average} = -0.001\ (\text{bio13} - \text{bio14}) + 0.265\ \text{bio15} + 0.184$$

The $r^2$ value for the above fit was found to be 0.81.

To make the sum of the monthly variances equal to the annual variance, the following scaling co-efficient is calculated,

$$c = \frac{\text{annual average} \times \text{annual variation}}{\sqrt{\sum (\text{monthly averages})^2}}$$

**Table 3**
Calculated bio-climatic variables from the historical weather data.

| City | Bio3 | Bio4 | Bio13–Bio14 | Bio15 |
|---|---|---|---|---|
| Beijing | 29.10% | 10.56 | 177.9 mm | 103.97% |
| London | 37.40% | 5.50 | 22.38 mm | 14.40% |
| Marseille | 34.70% | 6.78 | 45.95 mm | 36.47% |
| New York | 24.50% | 8.98 | 34.50 mm | 10.26% |

### 3.3. Stage 2: Crop suitability and risk calculation

There are four sequential steps in Stage 2.

*Step 2.1: Identifying optimal conditions for crop growth*

In the second stage, the suitability and risk of different geographical locations to grow certain crops are estimated taking into account the weather uncertainties calculated in the first stage. Suitability is used to describe a region's appropriateness to grow a particular crop as a factor of the region's atmospheric conditions. Risk is used to capture those situations when a region's suitability differs from its expected mean value more than a threshold which describes the risk appetite of the decision maker. This is done by running simulations that take into account different projections for future weather changes.

This step aims to identify the conditions affecting crop growth. Crops have distinct requirements in terms of various atmospheric conditions that will allow the crops to grow properly. This data can be obtained from the FAO Ecocrop database.[3] The database provides the suitable and optimal temperature (maximum and minimum) and average rainfall needed for a particular crop. Additional information that further influences crop growth, such as the altitude and soil type of a region can also be obtained from this database, and can be used to better determine a region's suitability. As mentioned in Stage 1, in this particular study, we focus only on the environmental conditions related to temperature and rainfall.

*Step 2.2: Selection of future climate data*

The aim of this step is to select the appropriate future climate data based on which future suitability will be measured. We use estimations provided by other researchers in the WORLDClim database.[4] Two decisions need to be made before using the WORLDClim database:

- Selection of greenhouse gas scenarios (Representative Concentration Pathways RCP)
  Future changes in climate are highly uncertain and depend on a range of variables including socio-economic factors, energy and land use, and emissions of greenhouse gases and other pollutants (van Vuuren et al., 2011). The future projects of greenhouse gas emissions and the dependent changes in weather are defined across four levels, called Representative Concentration Pathways (RCP) (IPCC, 2014). These four levels—RCP2.6, RCP4.5, RCP6.0 and RCP8.5—represent optimistic to extreme pessimistic views on greenhouse gas emissions. In this study we will work with RCP6.0, which corresponds to a peak in emissions around year 2080. However, the choice is arbitrary and the interested research can experiment with other RCPs.
- Selection of climate models (GCMs)
  Weather changes are generally related to physical, chemical, hydrological and oceanic processes and are subjected to greenhouse gas emissions (Taylor, Stouffer, & Meehl, 2012). Currently, there are

[3] http://ecocrop.fao.org/ecocrop/srv/en/home.
[4] http://www.worldclim.org/.

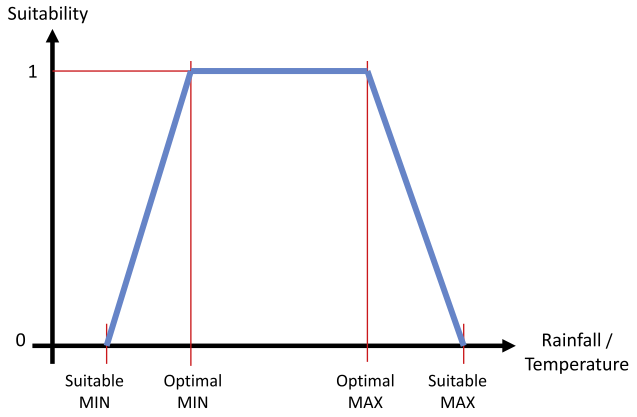


**Fig. 2.** Suitability calculation based on Ecocrop data.

many Global Circulation Models (GCMs) that are available which are based on different climate projection experiments, and realistic physics. The different changes in weather will influence the suitability of crops and it is important to choose a set of GCM models to calculate it. In this study we selected the following models and averaged our final results: (*a*) BCC-CSM1-1, (*b*) CCSM4, (*c*) GFDL-ESM2G, (*d*) HadGEM2-ES, (*e*) MRI-CGCM3, (*f*) NorESM1-M.

*Step 2.3: Definition of suitability and risk indices*

This step aims to define an way to measure the suitability and risk of a region in terms of growing a particular crop. This suitability index is defined based on both suitable and optimal temperature and rainfall ranges (obtained from the previous step), as follows:

$$Suitability = \begin{cases} 0 & \text{if } x < Suitable^{Min} \text{ or } x > Suitable^{Max} \\ \frac{x - Suitable^{Min}}{Optimal^{Min} - Suitable^{Min}} & \text{if } Suitable^{Min} \leqslant x \leqslant Optimal^{Min} \\ 1 & \text{if } Optimal^{Min} < x < Optimal^{Max} \\ \frac{x - Suitable^{Max}}{Optimal^{Max} - Suitable^{Max}} & \text{if } Optimal^{Max} \leqslant x \leqslant Suitable^{Max} \end{cases}$$

The above equation can be graphically depicted in Fig. 2.

As two different values for suitability can be measured (one based on the temperature and one based on rainfall), we use their product as the overall suitability index of a region.

We also define the notion of a risky region. A region is considered risky to grow a particular crop if the calculated suitability at any particular simulation iteration (see next step) is less than the mean value by a certain percentage. This indicates the risk appetite of the decision maker. For instance, a 10% threshold refers to the situation when a decision maker considers a region as being risky to grow a particular crop if the calculated suitability of that region is below 10% of the expected mean value. The risk index of a certain region is then measured as the percentage of times a region was found to be risky in our simulation study (see next step). Similar to suitability, two different values can be measured for risk (one based on the temperature and one based on rainfall). We consider the temperature-associated and the rainfall-associated risks to be independent. As a consequence, the overall risk index is calculated as:

$$Risk = 1 - (1 - risk_{temp}) * (1 - risk_{rainfall})$$

*Step 2.4: Calculation of suitability and risk*

The aim of this step is to calculate the suitability and risk for a given region and crop. This is done by firstly simulating various climate scenarios using the future climate data of step 2.2 and the weather uncertainty distributions described in Stage 1. At each iteration of the simulation (we ran 500 iterations in this study), a different set of values for temperature and rainfall is generated. These values are then used to calculate the suitability and risk of a region at each iteration (as described in step 2.3) and the mean value (of the 500 iterations) is taken as the suitability and risk of a particular region.

### 3.4. Stage 3: Modelling sourcing decisions

In the third stage, the estimated suitability and risk are used in a model that enables the optimisation of sourcing decisions taking into account both critical factors. This stage utilises the mean-variance model to decide what is the optimal way to split orders among different suppliers aiming to maximise profit while minimising variance.

The aim of this stage is to define a model to allow a customer to choose suppliers for a given crop on the basis of profit and variance. The sourcing decision model is modelled as a supplier selection problem, where suppliers are supplying regions where the food is sourced from. As an order of magnitude, the supplying regions in this study are millions of square kilometres wide. We model the sourcing decision as a mean-variance model, which takes a list of suppliers, their suitability and risk, and determines the quantity that should be ordered by each of them for a given demand scenario. The mean-variance model is used as it allows the representation of the risk appetite of the sourcing company, where there is a trade-off between maximising profit while minimising risk (Ray & Jenamani, 2016). In this paper, we assume the demand to be deterministic as it will be difficult to obtain the probability distribution of the demand of a particular crop in the future. The mean variance model is presented below:

The quantity delivered by a supplier can be written as:

$$m_i = Q_i \times y_i \times b_i + Q_i \times (1 - b_i)$$

where $Q_i$ is the quantity ordered to supplier $i$, $y_i$ represents which part of the order is met in case of failure, $b_i$ is a binary variable representing the success/failure of the supplier (1 = failure).

The profit for the sourcing company can be written as:

$$\Pi = \begin{cases} s \times \sum_{i=1}^{n} m_i - k \times (x - \sum_{i=1}^{n} m_i) - c \times \sum_{i=1}^{n} m_i & \text{if } \sum_{i=1}^{n} m_i \leqslant x \\ s \times x - r \times (\sum_{i=1}^{n} m_i - x) - c \times \sum_{i=1}^{n} m_i & \text{if } \sum_{i=1}^{n} m_i \geqslant x \end{cases} \quad (1)$$

where $s$ is the selling price for one unit, $r$ is the value of one unit of goods in inventory, $k$ is the loss of goodwill in the case of one unit of unmet demand, $x$ is the demand, $c_i$ is the cost of one unit from supplier $i$.

The above profit equation can be simplified as follows:

$$\Pi = \sum_{i=1}^{n} m_i \times (s - c_i + k) - k \times x - (s - r + k)(\sum_{i=1}^{n} m_i - x)\delta_{\sum_{i=1}^{n} m_i > x} \quad (2)$$

where $\delta$ represents the Kronecker delta (worth 1 if the condition associated is met, 0 otherwise).

From the above it is possible to derive the mean and variance as follows:

$$E(\Pi) = \sum_{i=1}^{n} \alpha_i E(m_i) - kx - \beta E(\int_0^{\Sigma m_i} \delta_{\sum_{i=1}^{m} m_i > x}(t)dt) \quad (3)$$

$$var(\Pi) = \sum_{i=1}^{n} \alpha_i^2 var(m_i) + \beta^2 var(\int_0^{\Sigma m_i} \delta_{\sum_{i=1}^{n} m_i > x}(t)dt) - 2\beta \sum_{i=1}^{n} \alpha_i cov(m_i, \int_0^{\Sigma m_i} \delta_{\sum_{i=1}^{m} m_i > x}(t)dt) \quad (4)$$

Then according to the mean-variance model, the maximising function can then be written as:

$$MV(\Pi) = E(\Pi) - A \times var(\Pi) \quad (5)$$

where $A$ is the risk averseness. This is the factor used in order to determine how much value a sourcing company is ready to trade for a certain decrease of variance (i.e. uncertainty).

Using the above model, it is possible to calculate the changes in value due to sourcing a crop from a particular region.

## 4. Case example

### 4.1. Introduction

In this section, we demonstrate the application of the approach by analysing the impact of climate on sourcing three crops: banana, cocoa beans and barley. As mentioned earlier, in this particular study, we use RCP60 as a representative indication of green house emissions and according to this selection, the green houses gases are expected to peak in the year 2080. Further, we compare the current suitability with the future in the year 2050. Different regions are selected for the selected crops depending on the current suitability.

### 4.2. Stage 1: Modelling climate uncertainty

We use the weather uncertainty model developed in Section 3.2. From that section, the temperature uncertainty is represented as a normal distribution whose variance is equal to the monthly variances. Gamma distribution is used to represent the uncertainty related to rainfall and its variance is such that the sum of the monthly variances equals the annual variance. As this stage does not depend on the specific crops under consideration, the model presented in Section 3.2 can be used as is.

### 4.3. Stage 2: Crop suitability and risk calculation

The impact of climate change varies depending on the growing regions and the crops under consideration. For example, an increase in temperature can lead to the shortening of the growing period and may lead to reduced yields. It can also change the area of cultivation either by turning current cultivated regions into unsuitable ones or by turning current uncultivated regions into more suitable ones. To understand these issues, the current and future climate scenarios for each crop are compared using with the temperature and rainfall thresholds obtained from Ecocrop. In this study, we look at various regions of interest, selected based on current cultivation of the chosen crops. Tables 4–6 summarise the calculated suitability and risk for each crop-area combination and the resulting size of the geographical area suitable for growing a particular crop. A risk threshold of 10% is used when calculating the risk index.

For *banana*, north and south of South America, central Africa and Oceania are considered. In Table 4, it can be seen that the total cultivated area of South America (North) decreases, whereas the area available for banana cultivation in central Africa and Oceania increases.

To assess the current and future suitability of *cocoa beans*, similar regions to the ones used for banana were considered. Here, the total cultivated area and the suitability of South America (North) decreases significantly. On the other hand, more regions become suited for cocoa cultivation in Central Africa (see Table 5).

In the case of *barley*, we compared South America, Western Europe, Eastern Europe, United States of America and Australia. The suitability decreases marginally in South America and United States (see Table 6). On the other hand, the suitability increases in Western and Eastern Europe along the cultivated area. Therefore, barley might be more suitable to be grown in Europe in the future. Based on this suitability in different regions, we assess the sourcing strategy in the next section.

**Table 4**
Current and future suitability of banana.

| | | South America (North) | South America (South) | Central Africa | Oceania |
|---|---|---|---|---|---|
| Current | Suitability (%) | 71.90 | 57.45 | 64.38 | 80.11 |
| | Risk (%) | 23.34 | 26.07 | 24.18 | 25.75 |
| | Area (km$^2$) | 4,457,100 | 1,125,500 | 1,125,500 | 3,442,100 |
| Future | Suitability (%) | 60.09 | 58.13 | 69.39 | 68.70 |
| | Risk (%) | 36.66 | 28.06 | 24.89 | 36.48 |
| | Area (km$^2$) | 3,069,800 | 1,298,9000 | 3,144,400 | 2,849,000 |

### 4.4. Stage 3: Modelling sourcing decisions

To decide on the optimal sourcing strategy, the following numerical values are used:

- The selling price per unit $s$ is £1.00.
- The value of one unit in inventory $r$ is assumed as 10% of selling price $s$, which is £0.1.
- The loss of good will in the case of unmet demand $k$ is assumed as 25% of selling price $s$, which is £0.25.

In this example, the cost is assumed as a function of suitability. The cost is taken as inversely proportional to the suitability. It costs less to produce in more suitable regions or it is expensive to produce in less suitable regions. The following function is used to represent the cost:

$$\frac{suitability}{50} = \frac{60}{cost}$$

which means at 100% suitability, it costs £0.3 to procure one unit.

We employ the mean-variance model to each combination of supplying regions to assess the profit and variance. The results obtained for various banana, cocoa and barley are shown in Tables 7–9.

For banana, the current optimal ordering strategy is to order mostly from north of South America. However, in the future, it will be optimal order mostly from central Africa, as this region becomes more suitable while south American suitability decreases. The current mean profit and variance is given as 36.60 and 12.38. The future mean and variance is given as 37.08 and 14.44, implying that the regions will be more suitable in the future for growing bananas, however, with increased variance (risk).

Similar to banana, the impact of climate change indicates that it will be more suitable to order from central Africa than from South America. The current mean profit decreases from 47.84 to 39.46 and the variance increases from 14.95 to 15.26. Therefore, cocoa beans have a higher risk, that is to meet a demand of 100 units, the current optimal order increases from 178 to 199. In this instance, companies might need to put more effort in improving the management of cocoa crops to minimise the risk of failure.

In the case of barley, the situation is fairly stable and it becomes preferable to source barley from Western Europe, as this region becomes more suitable. The mean profit and variance increases from 35.56 and 11.55 to 38.86 and 12.46 respectively. This implies that the climate change in general will have a positive impact on the barley supply chain by making more regions suitable for cultivation.

In general, the results suggest that areas of low suitability are becoming more favourable while the current high suitability areas seems to become less suited. Consequently, this means that the yields will go down and the crops might require more resources to produce. Due to climate change, companies might have to reconsider the region where they source the crops from or have to work with current suppliers either in mitigating or developing processes to overcome the risks attributed to climate change.

**Table 5**
Current and future suitability of cocoa beans.

| | | South America (North) | South America (South) | Central Africa | Oceania |
|---|---|---|---|---|---|
| Current | Suitability (%) | 77.84 | 59.16 | 65.27 | 75.24 |
| | Risk (%) | 23.97 | 33.85 | 29.93 | 25.74 |
| | Area ($km^2$) | 634,460 | 701,040 | 2,694,900 | 6,140,500 |
| Future | Suitability (%) | 68.62 | 57.05 | 67.18 | 67.41 |
| | Risk @10% (%) | 32.62 | 33.96 | 30.26 | 32.80 |
| | Area ($km^2$) | 542,250 | 830,310 | 2,899,700 | 5,683,200 |

**Table 6**
Current and future suitability of barley.

| | | South America | Western Europe | Eastern Europe | United States | Australia |
|---|---|---|---|---|---|---|
| Current | Suitability (%) | 65.54 | 62.55 | 61.20 | 62.96 | 65.72 |
| | Risk @10% (%) | 26.40 | 22.84 | 21.93 | 27.23 | 24.78 |
| | Area ($km^2$) | 2,797,400 | 1,127,500 | 1,129,600 | 1,637,500 | 1,534,400 |
| Future | Suitability (%) | 61.54 | 64.73 | 64.84 | 60.38 | 63.37 |
| | Risk @10% (%) | 26.62 | 23.55 | 22.08 | 27.85 | 25.42 |
| | Area ($km^2$) | 2,394,600 | 1,146,100 | 1,847,500 | 1,752,300 | 1,374,900 |

**Table 7**
Optimal Ordering Strategy of Banana for a demand of 100 units.

| Region | Current optimal ordering | Future optimal ordering |
|---|---|---|
| South American (North) | 127.58 | 10.41 |
| South America (South) | 0 | 8.71 |
| Central Africa | 3.64 | 117.03 |
| Oceania | 56.77 | 55.89 |

**Table 8**
Optimal Ordering Strategy of Cocoa-Beans for a demand of 100 units.

| Region | Current optimal ordering | Future optimal ordering |
|---|---|---|
| South American (North) | 100 | 59.37 |
| South America (South) | 5.66 | 8.01 |
| Central Africa | 11.46 | 115.11 |
| Oceania | 60.75 | 16.31 |

**Table 9**
Optimal Ordering Strategy of Barley for a demand of 100 units.

| Region | Current optimal ordering | Future optimal ordering |
|---|---|---|
| South America | 10.23 | 8.49 |
| Western Europe | 0.00 | 50.29 |
| Eastern Europe | 106.2 | 100.00 |
| United States | 8.64 | 0.00 |
| Australia | 49.54 | 9.96 |

## 5. Conclusions and future work

In this work, we used publicly available climate data and crop growth requirements to analyse the future impact of climate change on global food sourcing decisions. A generic approach was proposed for this analysis and was illustrated in a case example. In the case example, we presented the analysis for three different crops, banana, cocoa beans and barley. The analysis showed that some regions might become more suitable to grow certain crops in the future, while regions currently very suitable will become less so. This approach can be used by organisations to analyse a particular crop and region of interest to determine the optimal future sourcing.

In this paper, we developed and demonstrated an approach, where an exist sourcing decision model namely the mean-variance model, was adopted to include climate change data, especially in the context of crop production locations. The key feature of the model involves around big-data analytics. Here, we considered the data on geographically dispersed crop suitability data and associated climatic data variables falls under the remit of big-data. And the approach presented here scales with large and new datasets.

Mean-variance models have been used within the context of sourcing decisions under disruption risk. In this paper, we have extended this analysis for location decisions within the context of climate change. Additionally, this approach also contributes to adopting supplier selection decisions, dual sourcing and pricing strategies. Further, in this paper, we have used publicly available data from WorldClim and FAO to model the climate change impact on food sourcing decisions, thereby presenting an approach to utilise public domain data for organisational operational improvements.

One of the limitations of this study is that it focuses mainly on climate data to investigate the future of sourcing decisions. Future research should also include other important factors affecting crop growth such as soil type as well as non crop-related factors affecting sourcing decisions such as the political stability of a region. It is also important to examine how manufacturers and retailers using food items in their processes should prepare for the forthcoming changes by, for example, developing new relationships with suppliers from different geographical regions. Future work on improving the proposed approach can also explore applying more advanced data analytic techniques on the climate data to better model the weather uncertainties as the analysis relies heavily on weather predictions.

## Appendix A. Scoring model for weather uncertainties

See Table A.10.

**Table A.10**
Scoring system for modelling weather uncertainties.

| Score | Condition |
|---|---|
| 20 | $p \leqslant 90.00\%$ |
| 19 | $90\% < p \leqslant 95.00\%$ |
| 17 | $95\% < p \leqslant 97.50\%$ |
| 15 | $97.5\% < p \leqslant 99.00\%$ |
| 13 | $99.00\% < p \leqslant 99.50\%$ |
| 11 | $99.50\% < p \leqslant 99.90\%$ |
| 08 | $99.90\% < p \leqslant 99.95\%$ |
| 03 | $p > 99.95\%$ |

## Appendix B. Scores obtained for different temperature models

See Table B.11.

**Table B.11**
Scoring for temperature models.

| Temperature model | Min temp score | Max temp score | Total score |
|---|---|---|---|
| 1 | 965 | 886 | 1851 |
| 2 | 616 | 650 | 1266 |
| 3 | 843 | 912 | 1755 |
| 4 | 1078 | 1179 | 2257 |
| 5 | 867 | 887 | 1754 |
| 6 | 859 | 950 | 1809 |
| 7 | 911 | 1089 | 2000 |
| 8 | 834 | 853 | 1687 |

## Appendix C. Scoring obtained for different rainfall models

See Table C.12.

**Table C.12**
Scoring for rainfall models.

| Rainfall model | Normal distribution score | Gamma distribution score |
|---|---|---|
| 1 | 905 | 1025 |
| 2 | 858 | 1407 |
| 3 | 663 | 1073 |
| 4 | 913 | 1007 |
| 5 | 779 | 964 |
| 6 | 885 | 120 |
| 7 | 1038 | 132 |
| 8 | 714 | 201 |

## References


Addo-Tenkorang, R., & Helo, P. T. (2016). Big data applications in operations/supply-chain management: A literature review. *Computers & Industrial Engineering, 101*, 528–543.

Berger, P. D., Gerstenfeld, A., & Zeng, A. Z. (2004). How many suppliers are best? A decision-analysis approach. *Omega, 32*(1), 9–15.

Cogan, D. G. (2006). Corporate governance and climate change: Making the connection. Ceres Boston, MA.

Dada, M., Petruzzi, N. C., & Schwarz, L. B. (2007). A newsvendor's procurement problem when suppliers are unreliable. *Manufacturing & Service Operations Management, 9*(1), 9–32.

Demirtas, E. A., & Üstün, Ö. (2008). An integrated multiobjective decision making process for supplier selection and order allocation. *Omega, 36*(1), 76–90.

Federgruen, A., & Yang, N. (2008). Selecting a portfolio of suppliers under demand and supply risks. *Operations Research, 56*(4), 916–936.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management, 35*(2), 137–144.

Gunasekaran, A., Papadopoulos, T., Dubey, R., Wamba, S. F., Childe, S. J., Hazen, B., & Akter, S. (2017). Big data and predictive analytics for supply chain and organizational performance. *Journal of Business Research, 70*, 308–317.

Hammami, R., Temponi, C., & Frein, Y. (2014). A scenario-based stochastic model for supplier selection in global context with multiple buyers, currency fluctuation uncertainties, and price discounts. *European Journal of Operational Research, 233*(1), 159–170.

Heese, H. S. (2015). Single versus multiple sourcing and the evolution of bargaining positions. *Omega, 54*, 125–133.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology, 25*(15), 1965–1978.

Iceberg lettuces and broccoli rationed as vegetable crisis hits supermarkets (2017).

< https://www.bbc.co.uk/news/uk-38851097 > .
Intergovernmental Panel on Climate Change (IPCC) (2014). *Climate change 2014 – Impacts, adaptation and vulnerability: Regional aspects.* Cambridge University Press.
Kumar, M., Basu, P., & Avittathur, B. (2018). Pricing and sourcing strategies for competing retailers in supply chains under disruption risk. *European Journal of Operational Research, 265*(2), 533–543.
Linnenluecke, M., & Griffiths, A. (2010). Beyond adaptation: Resilience for business in light of climate change and weather extremes. *Business & Society, 49*(3), 477–511.
Meena, P., Sarmah, S., & Sarkar, A. (2011). Sourcing decisions under risks of catastrophic event disruptions. *Transportation Research Part E: Logistics and Transportation Review, 47*(6), 1058–1074.
Mehmood, R., Meriton, R., Graham, G., Hennelly, P., & Kumar, M. (2017). Exploring the influence of big data on city transport operations: A Markovian approach. *International Journal of Operations & Production Management, 37*(1), 75–104.
Mills, E., & Lecomte, E. (2006). From risk to opportunity: How insurers can proactively and profitably manage climate change. Ceres Report.
Patwardhan, A., Semenov, S., Schnieder, S., Burton, I., Magadza, C., Oppenheimer, M., ... Sukumar, R. (2007). Assessing key vulnerabilities and the risk from climate change. *Climate Change,* 779–810.
Ray, P., & Jenamani, M. (2016). Mean-variance analysis of sourcing decision under disruption risk. *European Journal of Operational Research, 250*(2), 679–689.
Ruiz-Torres, A. J., & Mahmoodi, F. (2007). The optimal number of suppliers considering the costs of individual supplier failures. *Omega, 35*(1), 104–115.
Sawik, T. (2011). Selection of supply portfolio under disruption risks. *Omega, 39*(2), 194–208.
Scheffer, M., Carpenter, S., Foley, J. A., Folke, C., & Walker, B. (2001). Catastrophic shifts in ecosystems. *Nature, 413*(6856), 591–596.
Snyder, L. V., Atan, Z., Peng, P., Rong, Y., Schmitt, A. J., & Sinsoysal, B. (2016). Or/ms models for supply chain disruptions: A review. *IIE Transactions, 48*(2), 89–109. https://doi.org/10.1080/0740817X.2015.1067735.
Sussman, F. G., & Freed, J. R. (2008). *Adapting to climate change: A business approach.* VA: Pew Center on Global Climate Change Arlington.
Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society, 93*(4), 485–498.
The tip of the iceberg: Climate change, lettuce shortages, and tesco's efforts to adapt its supply chain (2017). < https://rctom.hbs.org/submission/the-tip-of-the-iceberg-climate-change-lettuce-shortages-and-tescos-efforts-to-adapt-its-supply-chain/ > .
Tiwari, S., Wee, H., & Daryanto, Y. (2018). Big data analytics in supply chain management between 2010 and 2016: Insights to industries. *Computers & Industrial Engineering, 115*, 319–330.
van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., ... Rose, S. K. (2011). The representative concentration pathways: An overview. *Climatic Change, 109*(1), 5. https://doi.org/10.1007/s10584-011-0148-z.
Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics, 165*, 234–246.
Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M.-J. (2017). Big data in smart farming – A review. *Agricultural Systems, 153*, 69–80.
Wu, C., & Barnes, D. (2011). A literature review of decision-making models and approaches for partner selection in agile supply chains. *Journal of Purchasing and Supply Management, 17*(4), 256–274.
Yu, H., Zeng, A. Z., & Zhao, L. (2009). Single or dual sourcing: Decision-making in the presence of supply chain disruption risks. *Omega, 37*(4), 788–800.
Zhong, R. Y., Newman, S. T., Huang, G. Q., & Lan, S. (2016). Big data for supply chain management in the service and manufacturing sectors: Challenges, opportunities, and future perspectives. *Computers & Industrial Engineering, 101*, 572–591.